SIMULTANEOUS OBJECT CLASSIFICATION AND SEGMENTATION WITH HIGH-ORDER MULTIPLE SHAPE MODELS

By

Federico Lecumberry
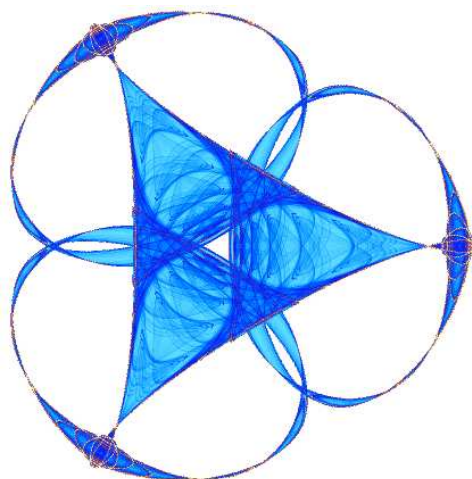
Álvaro Pardo

and

Guillermo Sapiro

INSTITUTE FOR MATHEMATICS AND ITS APPLICATIONS

Report Documentation Page

1. REPORT DATE **MAY 2009**	2. REPORT TYPE	3. DATES COVERED **00-00-2009 to 00-00-2009**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Simultaneous Object Classification and Segmentation with High-Order Multiple Shape Models	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of Minnesota,Institute for Mathematics and Its Applications,Minneapolis,MN,55455-0436**	8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
see report

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**11**	

Simultaneous Object Classification and Segmentation with High-Order Multiple Shape Models

Federico Lecumberry, *Student Member, IEEE,* Álvaro Pardo, *Member, IEEE,*
and Guillermo Sapiro, *Senior Member, IEEE*

Abstract—**Shape models (SMs), capturing the common features of a set of training shapes, represent a new incoming object based on its projection onto the corresponding model. Given a set of learned SMs representing different objects classes, and an image with a new shape, this work introduces a joint classification-segmentation framework with a twofold goal. First, to automatically select the SM that best represents the object, and second, to accurately segment the image taking into account both the image information and the features and variations learned from the on-line selected model. A new energy functional is introduced that simultaneously accomplishes both goals. Model selection is performed based on a shape similarity measure, on-line determining which model to use at each iteration of the steepest descent minimization, allowing for model switching and adaptation to the data. High-order SMs are used in order to deal with very similar object classes and natural variability within them. Position and transformation invariance is included as part of the modeling as well. The presentation of the framework is complemented with examples for the difficult task of simultaneously classifying and segmenting closely related shapes, such as stages of human activities, in images with severe occlusions.**

Index Terms—**Shape priors, image segmentation, object modeling, variational formulations.**

I. INTRODUCTION

Object segmentation is one of the most fundamental tasks in image processing, still lacking a completely automatic solution. The main idea is to find a set of features that describes and discriminates the object of interest from the rest of the image. Object color is a low level feature that can be used as such descriptor, although its discrimination capacity is often insufficient in real images. Using shape as a high level feature is a common approach to augment such low level features.

The shape of the desired object is added as a descriptor, constraining the set of possible solutions to regions of the image that simultaneously "match" this shape and the low level features (color, edges, etc.). The most common way to add this shape information is in the form of a weighted linear

F. Lecumberry is with the Instituto de Ingeniería Eléctrica, Facultad de Ingeniería, Universidad de la República, Montevideo, Uruguay. e-mail: fefo@fing.edu.uy. FL performed this work at the University of Minnesota.

Á. Pardo is with the Departamento de Ingeniería Eléctrica Facultad de Ingeniería y Tecnologías, Universidad Católica del Uruguay, Montevideo, Uruguay. e-mail: apardo@ucu.edu.uy.

G. Sapiro is with the Department of Electrical and Computer Engineering University of Minnesota, Minneapolis, MN, 55414 USA e-mail: guille@umn.edu.

combination of functionals addressing, on one hand, the low level features and, on the other hand, the shape priors or models. This leads to a minimization problem where the solution is a compromise between the shape of the final contour and the information given by the image. The minimization techniques used in the literature, include, among others, gradient descent methods [1]–[5] and graph-cuts [6]. The used shape representations can be signed distance functions (SDF) [1]–[3], [5], [6], quadratic splines [7], characteristic functions [4], and landmark points [8].

When M different objects (or object classes) can appear in an image, a single shape prior (model) is not sufficient, and multiple shape priors must be considered. A possible, but not elegant, approach is to run the process with each one of the shape priors separately, and then choose the best solution. Vu and Manjunath [6] and Cremers et al. [5] define M possible labels for each pixel on the image, and then propose a segmentation energy that includes the optimization of these labels in order to determine where to apply each prior. In a different work, Cremers et al. [7] perform density estimation in a non-linear feature space, where different objects are separable. The proposed energy is then minimized considering both the curve's control points and the image.

Considering the natural deformations and the variability of objects within a class, high-order shape models (SMs) should be included in the segmentation. Leventon et al. [1] compute PCA on a set of registered shapes, fitting a Gaussian probability distribution to the coefficients of the reconstruction. This allows to include the probability of a certain shape, in traditional geodesic active contours for low level features, and a MAP estimation of the object in the image. Tsai et al. [3] also use PCA to model shape variations, defining an energy for the aligning of the binary shape, and formulate a segmentation functional optimizing the parameters of the representation with the first deformation modes. Cootes and Taylor [8] compute, using PCA, a point distribution model of landmarks points defining a shape. More recently, Charpiat et al. [4] proposed a framework to compute non-linear shape statistics based on the Hausdorff distance between shapes, and then model distributions similarly to [1].

In this work, a new framework for image segmentation with multiple high-order shape models is introduced, addressing at the same time the selection of the model and its image-driven positioning and adjustment to the modeled deforma-

(a) Shapes from class 1   (b) Modes of variation for class 1   (c) Shape (left) and its five projections onto five models

(d) Modes of variation for class 2   (e) Modes of variation for class 3   (f) Modes of variation for class 4   (g) Modes of variation for class 5
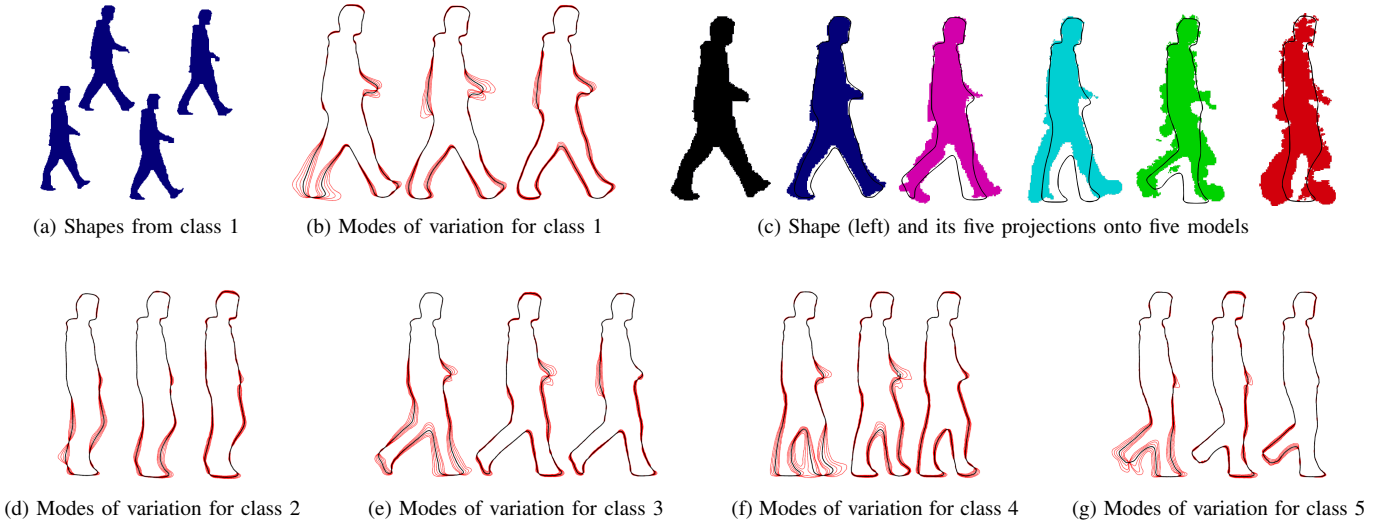
Fig. 1. (a) Four shapes from one of the classes, and (b) the first three modes of variation of the corresponding model in the walking sequence. The thick black line is the mean shape, the red lines are obtained varying the amplitude (see text). (c) Original shape (in black) and its projections $\mathcal{P}_k\phi$ (colored as in Figure 2) to $M = 5$ different models in the walking sequence, one for each cluster. The mean shape of the corresponding model is plotted too (black curve). The projections are ordered based in the measure given by Equation (8). Note how the projection is completely deformed when using the wrong shape model. (d) - (g) First three modes of variation for four different shape models. (This figure is in colors.)

tions. Invariance is included as part of the framework as well. In particular, the high-order SMs are computed using PCA in a similar way as [1], [3], obtaining a set of eigenmodes of variations. In the case of dynamic shapes with large, non-linear deformations, a method to obtain a lineal approximation of the shape space is described using a dimensionality reduction algorithm. The selection of the model is obtained with a binary selection coefficient, on-line learned based on a similarity measure between shapes. The proposed framework follows from a functional that combines two terms. The first one is a region-based segmentation term [9]. The second term is a combination of the multiple high-order SMs, addressing the model selection and constraining the solution to the high-order shape information coming from the on-line selected model. While the framework is presented for planar curves, it can be easily extended to data in higher dimensions.

The remainder of the paper is organized as following. Section II reviews briefly the definition and properties of shapes models. Section III describes the proposed framework. Section IV presents experiments testing the ideas and theirs discussion. Section V proposes an invariance to translation extension of the framework. Finally, Section VI concludes the work.

II. HIGH-ORDER MULTIPLE SHAPES MODELS

Consider M sets Φ_k, $k = 1, \ldots, M$, each with N_k registered shapes $\Phi_k = \{\phi_k^1, \ldots, \phi_k^{N_k}\}$, where each ϕ_k^j is a signed distance function (SDF), whose zero level-set curve, $\mathcal{C}_k^j$, represents a shape from the k-th class of objects. Let $\mathcal{M}_k^d$ be a d-order model that captures the intrinsic deformations of the training set Φ_k for the class k. In this work, $\mathcal{M}_k^d$ is derived from a PCA decomposition of the training set Φ_k (all the shapes ϕ_k are represented as vectors in $\mathbb{R}^D$, D being the size of the range of the corresponding SDFs),

$$\mathcal{M}_k^d := \{\mu_k, U_k^d\}, \tag{1}$$

where $\mu_k \in \mathbb{R}^D$ is the mean shape of Φ_k, $U_k^d \in \mathbb{R}^{D \times d}$ is a matrix containing the first d modes of variation (eigenmodes), $U_k^d = [\{u_k^i\}_{i=1}^d], u_k^i \in \mathbb{R}^D$.

A model $\mathcal{M}_k^d$ generates a representation of a new incoming shape $\hat{\phi}$ by the d-projection $\mathcal{P}_k^d \hat{\phi}$

$$\mathcal{P}_k^d \hat{\phi} = \mu_k + U_k^d \alpha_k, \tag{2}$$

where $\alpha_k \in \mathbb{R}^d$ are the corresponding reconstruction coefficients, which of course depend on $\hat{\phi}$ (see for example [1, section 2.1] for details).

The accuracy of the representation depends on the similarity between $\hat{\phi}$ and the shapes in Φ_k. Constraining small shape variations in the class Φ_k (compared with the deformations across different classes k) allows to obtain accurate representations using a linear approximation like PCA.

Finally, let

$$\mathfrak{M} = \{\mathcal{M}_1, \ldots, \mathcal{M}_M\}$$

be a set of SMs for the M different classes of objects. For simplicity, the order d in the notation is omitted from now on.

Figure 1 shows SMs for a walking person. Figure 1a shows four different shapes from one of the classes of shapes, note the similarity between them. Figure 1b shows the first three modes of variation of the corresponding model in a walking sequence. The data was obtained filming a single person walking with a static background [10]. The thick black line is the zero level set of the mean shape. The red lines are the zero level sets of the addition of the mean shape and a constant times the first, second, or third eigenmode respectively, varying the amplitude. Figure 1c shows an original shape from the set and its projections (with $d = 21$) to $M = 5$ different models in a walking sequence. The mean shape of the corresponding model is plotted too (black curve). The projections are ordered based in the measure given later by Equation (8). Note how the projection is completely deformed when using the wrong model, clearly illustrating the importance of selecting the
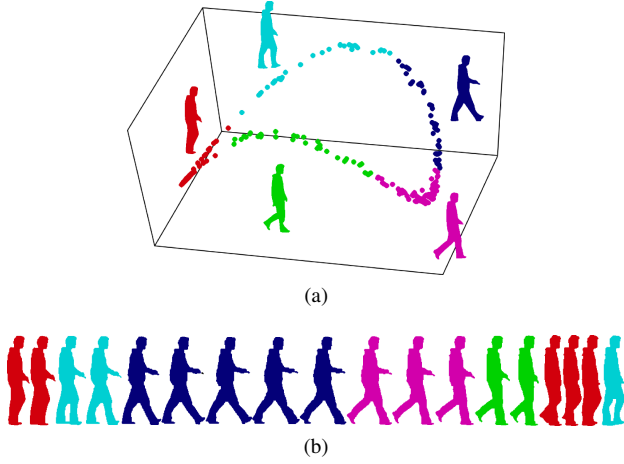
Fig. 2. (a) Low dimensional embedding. Each point correspond to the first three coordinates of the mapping obtained with DM. The colors correspond to the $M = 5$ obtained cluster. One sample shape from each cluster (walking-cycle position) is shown. (b) Eighteen samples from the walking sequence colored based on the obtained clusters. (This figure is in colors.)

variations already well approximated by ordinary PCA.

Figure 2 shows the clustering result. Figure 2a shows the low dimensional embedding manifold. Each point correspond to the first three coordinates of the mapping colored based on the obtained clusters. One sample shape from each cluster (walking-cycle position) is shown too. Figure 2b shows eighteen consecutive samples from the walking sequence colored based on the obtained clusters.

III. PROPOSED VARIATIONAL FRAMEWORK

Given an input image $\mathcal{I} : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ containing one or more shapes generated by the shape models $\mathcal{M}_k \in \mathfrak{M}$, an energy E is defined to simultaneously select the best model(s) and obtain a segmentation of the corresponding objects in $\mathcal{I}$ (a single object in each image is considered from now on for simplicity),

$$(\mathcal{M}^*, \phi^*) := \arg \min_{\substack{\mathcal{M} \in \mathfrak{M}, \\ \phi, c_+, c_-}} E(\mathcal{I}, \phi, c_+, c_-, \mathfrak{M}). \quad (3)$$

This energy includes two terms linearly combined with the constant λ,

$$E(\mathcal{I}, \phi, c_+, c_-, \mathfrak{M}) = E_{\text{CV}}(\mathcal{I}, \phi, c_+, c_-) + \lambda E_{\text{SM}}(\phi, \mathfrak{M}). \quad (4)$$

The E_{CV} term is, for the examples in this paper, the energy introduced by Chan and Vese [9],

$$\begin{aligned} E_{\text{CV}}(\mathcal{I}, \phi, c_+, c_-) &= \int_{\Omega} |\mathcal{I}(x) - c_+|^2 H(\phi(x)) \mathrm{d}x + \\ & \int_{\Omega} |\mathcal{I}(x) - c_-|^2 (1 - H(\phi(x))) \mathrm{d}x + \\ & \mu \int_{\Omega} \delta(\phi(x)) |\nabla \phi(x)| \mathrm{d}x, \quad (5) \end{aligned}$$

where c_+ and c_- are the averages of the input data inside and outside the curve $\mathcal{C}$ (the zero level set of ϕ), respectively, $H(\cdot)$ is the Heaviside function, and $\delta(\cdot)$ is the Dirac function. This energy attempts to split the input data into two different regions of approximately piecewise constant color or gray level values (c_+ and c_-). Other low level descriptors could be used for a better discrimination, for example texture [15] or edges [16].

The term E_{SM} adds an additional force aiming at maximizing the similarity between the evolving shape ϕ and its projection onto *only one* of the d-order models from $\mathfrak{M}$. Which one of the M models is used depends on the evolution of the shape and its projection to each model. The proposed term is

$$E_{\text{SM}}(\phi, \mathfrak{M}) = \sum_{k=1}^{M} \beta_k E_k(\phi, \mathcal{M}_k), \quad (6)$$

defining

$$E_k(\phi, \mathcal{M}_k) = \int_{\Omega} \|H(\phi(p)) - H(\mathcal{P}_k \phi(p))\|^2 \mathrm{d}p, \quad (7)$$

where again $H(\cdot)$ is the Heaviside function, β_k is a binary coefficient that (on-line) selects which of the M models is used, and $\mathcal{P}_k \phi$ is the projection of ϕ onto the model $\mathcal{M}_k$, given by Equation (2). Only one of the β_k must be different from zero in (6), since it is not fair to penalize for models that do not correspond to the object in the image. This is detailed next.

correct shape model (prior). Figures 1d, 1e, 1f and 1g show the first three modes of variation of the other four models obtained from the same walking sequence. The procedure to obtain these models is explained in the next section.

A. Clustering a set of shapes

One of the used datasets of shapes was taken from the sequence of a walking person [10]. Considered as a unique deformable object, this shape has large, non-linear deformations, invalidating the hypothesis of small (and linear) shape variations for this set. To alleviate this, a set of clusters may be considered. In this way linear approximations can be used to approximate shape deformations within each cluster to obtain $\mathcal{M}_k$. In order to obtain the clusters, in this work a non-linear mapping to an Euclidean space is performed based on Diffusion Maps (DM) [11]. DM is a general framework for data analysis based on a diffusion process over an undirected weighted graph, defining a new metric on the data called Diffusion Distance. Two properties of this metric are important in the present work. First, as a consequence of the density renormalized kernel defined to build the graph, the graph-Laplacian (see VonLuxburg's tutorial [12] for definition and properties of the graph Laplacian) is an approximation of the Laplace-Beltrami operator on the underlying manifold, allowing to recover the Riemannian geometry of the data set regardless the distribution of the points in the underlying manifold. Second, the Diffusion Distance is equivalent to the Euclidean distance in the space with coordinates given by the mapping function. This allows to simply compute K-means in the corresponding Euclidean space in order to group the shapes into M clusters and then obtain a local model in each cluster.

To recapitulate, the clusters are obtained by mapping into a new space via DM (a kernel method) and then applying K-means on this space. Note that the subsequent PCA could actually be performed in this space as well (using Kernel PCA [13], [14]), though the clustering makes the inner class
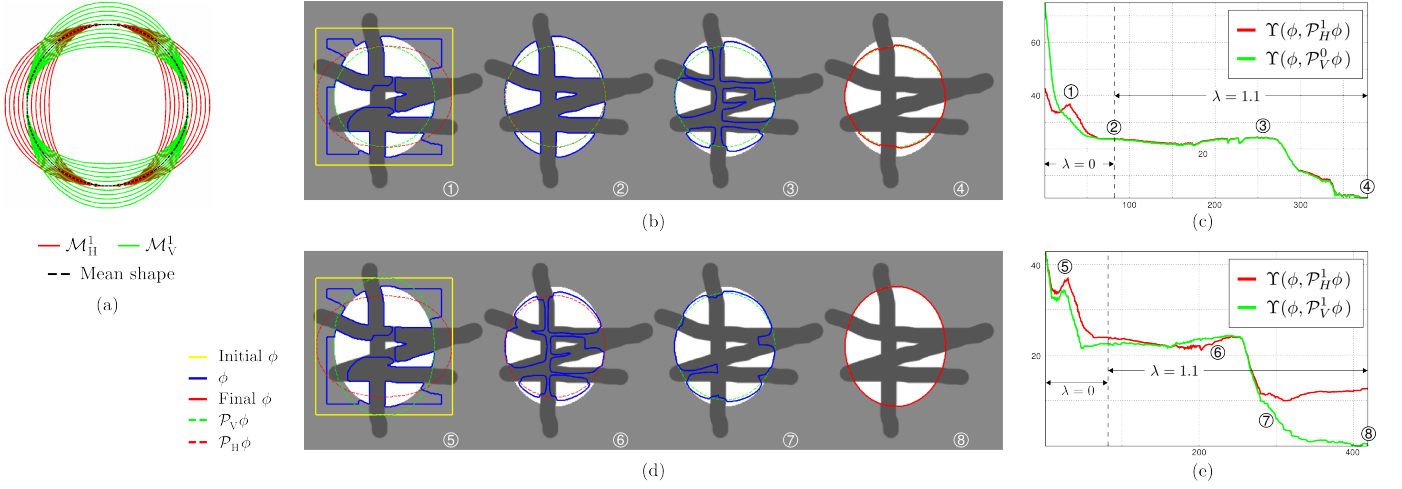
Fig. 3. (a) Mode of variation for the two ellipses models ($\mathcal{M}_\text{V}^1$ in green and $\mathcal{M}_\text{H}^1$ in red), the mean shape of both models is the same and is plotted in black dash line. (b) Results for the experiments with $\mathcal{M}_\text{H}^1$ and $\mathcal{M}_\text{V}^0$ (only mean shape). Some steps in the segmentation (see text) (c) Evolution of the shape dissimilarity measure for the experiments with $\mathcal{M}_\text{H}^1$ and $\mathcal{M}_\text{V}^0$. (d) Results for experiments with $\mathcal{M}_\text{H}^1$ and $\mathcal{M}_\text{V}^1$ (complete model). (e) Evolution of the shape dissimilarity measure for the experiments with $\mathcal{M}_\text{H}^1$ and $\mathcal{M}_\text{V}^1$. (This figure is in colors.)

A. Shape dissimilarity measure and model selection

Which is the non-zero β_k in Equation (6) is computed based on a shape dissimilarity measure (Υ) between two shapes ϕ_1 and ϕ_2,

$$\Upsilon(\phi_1,\phi_2) = \int_\Omega \frac{|\phi_1(p)|\delta(\phi_2(p))}{\text{length}(\mathcal{C}_2)}\mathrm{d}p + \int_\Omega \frac{|\phi_2(p)|\delta(\phi_1(p))}{\text{length}(\mathcal{C}_1)}\mathrm{d}p. \tag{8}$$

This is a length-normalized variation of the measure introduced by Funkhouser et al. [17]. This measure evaluates the sum of Euclidean distances corresponding to moving the contour of the first shape to points in the contour of the second shape, and viceversa, scaled by the curves lengths. In Figure 1c, the projected shapes are ordered according to increasing values of $\Upsilon(\phi,\mathcal{P}_k\phi)$. These ordered values are 1.35, 2.83, 3.59, 5.87, and 7.83 respectively.

Based on (8), a normalized shape similarity measure $\bar{\xi}_k(\phi)$ between a shape ϕ and its projection $\mathcal{P}_k\phi$ to the d-order k-th model is computed as

$$\bar{\xi}_k(\phi) = \frac{\xi_k(\phi)}{\sum_{l=1}^M \xi_l(\phi)}, \quad \text{where} \quad \xi_k(\phi) = \exp\left(-\Upsilon(\phi,\mathcal{P}_k\phi)\right). \tag{9}$$

This normalized similarity measure $\bar{\xi}_k(\phi)$ is close to one for the model that better represents ϕ. Finally to force the binary value in β_k, soft thresholding, based on a sigmoid function, is performed. Note that a unique coefficient is used as model selector, instead of one coefficient in each pixel as in [5], [6]. This encourages shape consistency and significantly simplifies the optimization.

With the proposed method, one model is always selected and a segmentation is obtained, even if the shape in the image has no appropriate model in $\mathfrak{M}$ that provides a good representation. The validation of the final segmentation can not be directly compared to the original non-occluded shape in all the cases, since there is no way to "create" the particular features or attributes that are occluded in it. Instead of this, the resulting segmentation is evaluated taking into account

the fact that the shape is generated by a model $\mathcal{M}$ in $\mathfrak{M}$ and the solution should then be a "valid" shape generated by this model. The following measure permits to discard a segmentation $\tilde{\phi}$ given the selected model. First, the mean $\bar{\Upsilon}_k$ and variance $\sigma^2_{\Upsilon_k}$ of $\Upsilon(\phi_k^j,\mu_k)$ are computed $\forall \phi_k^j \in \Phi_k$. Then if

$$\Upsilon(\tilde{\phi},\mu_k) > \bar{\Upsilon}_k + 1.5\,\sigma^2_{\Upsilon_k}, \tag{10}$$

the segmentation is discarded, and the shape can not be recognized.

B. Energy minimization

The proposed energy is minimized using a classical gradient descent method. For the gradient descent of E_CV, the expression is given in [9, Equation (9)]

$$\frac{\partial E_\text{CV}}{\partial \phi} = \delta(\phi)\left((\mathcal{I}-c_+)^2 - (\mathcal{I}-c_-)^2 - \mu\,\nabla\cdot\left(\frac{\nabla\phi}{|\nabla\phi|}\right)\right).$$

For E_SM, the obtained expression is

$$\frac{\partial E_\text{SM}}{\partial \phi} = -2\sum_{k=1}^M \beta_k \left\|H(\phi) - H(\mathcal{P}_k\phi)\right\|\Big(\delta(\phi) - \delta(\mathcal{P}_k\phi)W\Big),$$

where $W = U_k U_k^T$. Although the model selector β_k depends of ϕ, is treated as static, as a first order approximation for the gradient descent, since it affects the model selection and only indirectly the evolution of the curve.

Finally the first variation of Equation (4) becomes,

$$\frac{\partial E}{\partial \phi} = \frac{\partial E_\text{CV}}{\partial \phi} + \lambda\frac{\partial E_\text{SM}}{\partial \phi}. \tag{11}$$

C. Prior activation

The first steps of the optimization are performed without SMs information ($\lambda = 0$), until stationarity, then the "prior is activated" adding E_SM with $\lambda \neq 0$ (manually selected) until a new stationary point is reached, now combining the image and the shape information. This helps to determine the object

in the image without affecting the initial steps of the evolution with the projections to the models of the initial curve used in the minimization, which in the general has no similarity with the shapes in the models. A similar idea of "prior activation" is considered by Vu and Manjunath [6] using shape prior templates instead of SMs.

The Gestalt Principles [18] can give some intuition to this initial step. The "Principle of Similarity" states that people try to organize visual elements into groups based in the similarity of certain feature (shape, intensity, texture, etc.). This gives an additional argument for trying to start grouping regions of similar intensity and use the results as an initial point or "primary units" for helping the minimization process. After the identification of these "primary units," the addition of priors is used for a better interpretation of the object or scene.

IV. EXPERIMENTAL RESULTS

A. Models of ellipses

The first example is a "toy example," though illustrative and challenging, with two models of ellipses, where the only difference is that the first (and only beyond the mean shape) eigenmode is rotated $\frac{\pi}{2}$ (this already exemplifies the importance of high-order models). Let us name $\mathcal{M}_{\mathrm{V}}^1$ the model with vertical deformations and $\mathcal{M}_{\mathrm{H}}^1$ the model with horizontal deformations. Figure 3a shows the mode of variation for both models, in green for $\mathcal{M}_{\mathrm{H}}^1$ and in red for $\mathcal{M}_{\mathrm{H}}^1$.

The input image contains an occluded vertical ellipse, not present in the training set. Two different experiments are presented, varying the order d of the model $\mathcal{M}_{\mathrm{V}}^d$ while maintaining the highest dimension for the model that does not represent the input shape, $\mathcal{M}_{\mathrm{H}}^1$. With $d = 0$, only the mean shape is considered in the shape prior (no deformations), with $d = 1$ the vertical deformations are considered. All the parameters are the same in both experiments. Figures 3b and 3d show some steps in the minimization, and figures 3c and 3d show the evolution of the shape dissimilarity measure, for both experiments, respectively. Steps ① and ⑤ show an intermediate curve in the evolution with $\lambda = 0$, and the projections, $\mathcal{P}_{\mathrm{H}}^1\phi$ and $\mathcal{P}_{\mathrm{V}}^d\phi$, to both models, dashed colored lines. The initial curve (in yellow) is also shown. Note that $\mathcal{P}_{\mathrm{V}}^0\phi$ has no vertical deformations. The following steps (②,③ and ⑥,⑦) show the evolution after the "prior activation" adding the E_{SM} term ($\lambda = 1.1$), and the obtained segmentation (④ and ⑧).

In the first experiment (Figure 3b), the projections to both models end in the same shape, the mean shape. This is reflected also in the graph of dissimilarity measure (Figure 3c) by the overlapping of the green and red curves. In the second experiment (Figure 3d), $\mathcal{M}_{\mathrm{V}}^1$ captures the variation of the input shape, as reflected in the obtained segmentation. In this case there is also a model switching around iteration 200 (step ⑥), where the $\mathcal{M}_{\mathrm{H}}^1$ is selected while the occlusions are being filled. After this point, the vertical deformation determines the selection of $\mathcal{M}_{\mathrm{V}}^1$ for the rest of the evolution, ending with an accurate segmentation. Clearly, the high-order model and the automatic model selection are critical to obtain the correct segmentation.
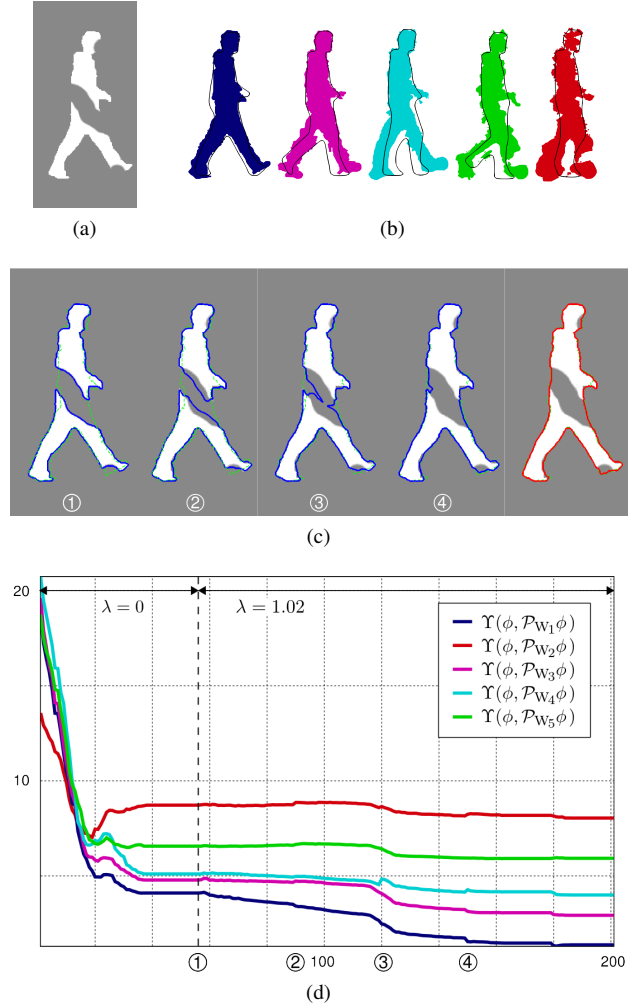

Fig. 4. (a) Input image with an occluded shape $\hat{\phi}_1$ in gray levels. (b) Projections of ϕ in the "prior activation" iteration (blue curve in step ① in Figure 4c) onto the five models, ordered based on $\Upsilon(\phi, \mathcal{P}_{\mathrm{W}_k}\phi)$. The mean shape of the corresponding model is plotted too (black curve). (c) Steps ① to ④ in the evolution of ϕ (blue curve) and its projections onto the selected model $\mathcal{M}_{\mathrm{W}_1}$ (green curve). The obtained segmentation is the red curve. (d) Evolution of the shape dissimilarity measure, $\Upsilon(\phi, \mathcal{P}_{\mathrm{W}_k}\phi)$ with the iterations. The curves in steps ①-④ are shown in Figure 4c. (This figure is in colors.)

B. Models from the walking sequence

Five high-dimensional models of a walking person cycle $\mathcal{M}_{\mathrm{W}_k}^{21}$ ($k = 1, \ldots, 5$, $d = 21$) were obtained with the procedure explained in Section II-A. The first three modes of variation for each model are shown in Figure 1. These are the models in the set of models $\mathfrak{M} = \{\mathcal{M}_{\mathrm{W}_k}\}_{k=1}^5$ for the next experiment. This set of models is particularly challenging for model selection since they are different deformations of the "same object."

The input image in this experiment contains a new occluded shape $\hat{\phi}_1$ (Figure 4a) that belongs to the model $\mathcal{M}_{\mathrm{W}_1}$ and is not in its training set Φ_1. Figure 4 shows details about the segmentation of $\hat{\phi}_1$. Figure 4c shows four steps after the "prior activation" (steps ① to ④) in the evolution of ϕ (blue curve) and their projections onto the automatically on-line selected model $\mathcal{M}_{\mathrm{W}_1}$ (green curve). Also the obtained segmentation $\tilde{\phi}_1$ (red curve) and its projection is shown. Figure 4b shows the projections of ϕ (blue curve) in the "prior activation" iteration onto the five models, ordered based on $\Upsilon(\phi, \mathcal{P}_{\mathrm{W}_k}\phi)$ for this

iteration. The mean shape of the corresponding model is plotted too (black curve). Compare the projections of the occluded shapes (Figure 4b) with those of a similar non-occluded shape in Figure 1c. Note how the projections onto the incorrect models are not too different, but the projections onto the correct model have significative differences. Figure 4d plots the evolution of the shape dissimilarity measure, $\Upsilon(\phi, \mathcal{P}_{w_k}\phi)$, for all the iterations and the five models. Note how the correct model is the one with lowest dissimilarity measure.

This experiment is repeated four times, maintaining the same set of models $\mathfrak{M} = \{\mathcal{M}_{w_k}\}_{k=1}^{5}$ and changing the input image. In each repetition, the input image contains a new occluded shape $\hat{\phi}_k$, $k = 2, 3, 4, 5$, belonging to the models $\mathcal{M}_{w_k}$, $k = 2, 3, 4, 5$, respectively. These four images are shown in Figure 5a with the corresponding obtained segmentations $\tilde{\phi}_k$ (red curves) and the projections onto the corresponding selected model (dashed green curves).

The results in Figure 4 show a number of important characteristics of the proposed framework that are consistent for all the presented experiments. First, in all the examples the selected model is the one to which the input shape belongs and the obtained segmentation is accurate to the data given by the image. Also, for all the experiments, during the minimization iterations the model selection is stable and there is no switch between the models once the shape prior is activated.

Second, it is relatively easy to follow the variations of the projection in the shape dissimilarity measure graph as the curve evolves. When the occlusions are being filled, the projection gets more similar to the shapes in $\mathcal{M}_{w_1}$ and the dissimilarity $\Upsilon(\phi, \mathcal{P}_{w_1}\phi)$ reduces. This is due to the force generated by the shape term E_{SM}, and as the curve gets closer to its projection this term attracts the curve strongly. Although this behavior is due to the E_{SM} term, the competition of both energy terms in areas of the shape where there is no occlusion preserves the curve close to the contour of the original shape, preventing to locally follow the projection, meaning that the E_{CV} energy term is stronger than the prior in this area of the image. This can be seen in the final segmentation (red curve in Figure 4c), where the projection (dashed green curve) in the hand goes through the original shape but the curve respects the gray level information. Similar details can be seen in Figure 5a for a different example. This example shows how the two energy terms collaborate to obtain a good segmentation of the occluded shape and each term attempts to define the curve in the regions where it better describes the solution. Where there is an occlusion the shape prior term takes control of
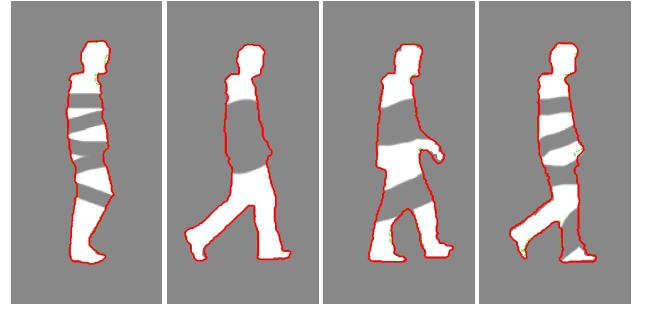
(a)
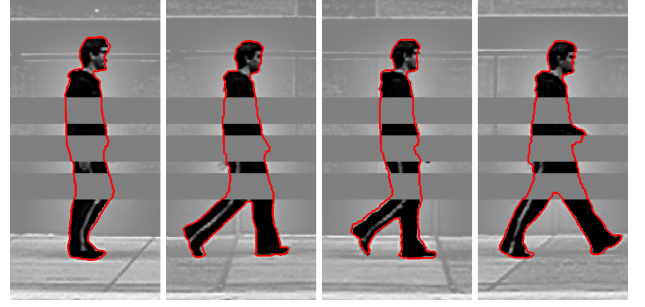
(b)

Fig. 5. Segmentations obtained with the proposed framework with the set of models $\mathfrak{M} = \{\mathcal{M}_{w_k}\}_{k=1}^{5}$ for different input images. (a) Segmentations for the binary occluded shapes $\hat{\phi}_2$, $\hat{\phi}_3$, $\hat{\phi}_4$ and $\hat{\phi}_5$ belonging to different models $\mathcal{M}_{w_k}$, $k = 2, 3, 4, 5$, respectively. (b) Segmentations of the gray level images with added occlusions. The shapes in these images also belong to different models $\mathcal{M}_{w_k}$, which are all correctly selected by the proposed framework. (This figure is in colors.)

the curve and where there is information of the actual shape (given by the intensity of the pixels), the data term controls the curve. This is done in a collaborative way, there is no discontinuity in the curve and it remains smooth. In order to achieve this, the projection onto the proper model is critical. Also the selection of the parameter λ is important, determining these collaboration/competition between both energy terms. In this work, as often done in the literature, λ is manually obtained. As a rule of thumb $\lambda \in (1.0, 1.2)$ was found to be a good initial estimation.

Table I shows numerical results for the validation of the obtained segmentation (see Equation (10)). The dissimilarity between the obtained shape and the original non-occluded shape $\Upsilon(\tilde{\phi}_k, \phi_k)$ (possible in these experiments since it is accessible) is shown in the last column. Note that these last measures are, in general, significantly smaller than the mean

$k(\mathcal{M}_{w_k})$	$\Upsilon(\tilde{\phi}_k, \mu_k)$	$\bar{\Upsilon}$	σ_{Υ}^2	$\Upsilon(\tilde{\phi}_k, \phi_k)$
1	4.53	3.85	0.91	0.86
2	3.39	2.66	0.68	1.05
3	2.60	3.42	1.31	0.59
4	4.14	3.75	1.37	1.44
5	2.99	3.18	1.17	2.18

TABLE I
NUMERICAL VALIDATION RESULTS (EQUATION (10)) AND DISSIMILARITY MEASURES (EQUATION (8)) FOR THE OBTAINED SEGMENTATIONS $\tilde{\phi}_k$ FOR THE EXPERIMENT IN FIGURES 4 AND 5.

	$\Upsilon(\tilde{\phi}_k, \mu_1)$	$\Upsilon(\tilde{\phi}_k, \mu_2)$	$\Upsilon(\tilde{\phi}_k, \mu_3)$	$\Upsilon(\tilde{\phi}_k, \mu_4)$	$\Upsilon(\tilde{\phi}_k, \mu_5)$
$\tilde{\phi}_1$	**4.53**	13.16	6.99	9.08	11.54
$\tilde{\phi}_2$	12.99	**3.39**	11.05	7.65	6.05
$\tilde{\phi}_3$	6.87	9.14	**2.60**	6.34	7.45
$\tilde{\phi}_4$	6.89	8.43	7.48	**4.14**	7.97
$\tilde{\phi}_5$	8.81	6.64	7.01	6.54	**2.99**

TABLE II
DISSIMILARITIES MEASURES (EQUATION (8)) BETWEEN THE OBTAINED SEGMENTATIONS $\tilde{\phi}_k$ AND THE MEAN SHAPES μ_k OF THE MODELS IN $\mathfrak{M}$ FOR THE EXPERIMENT IN FIGURES 4 AND 5.
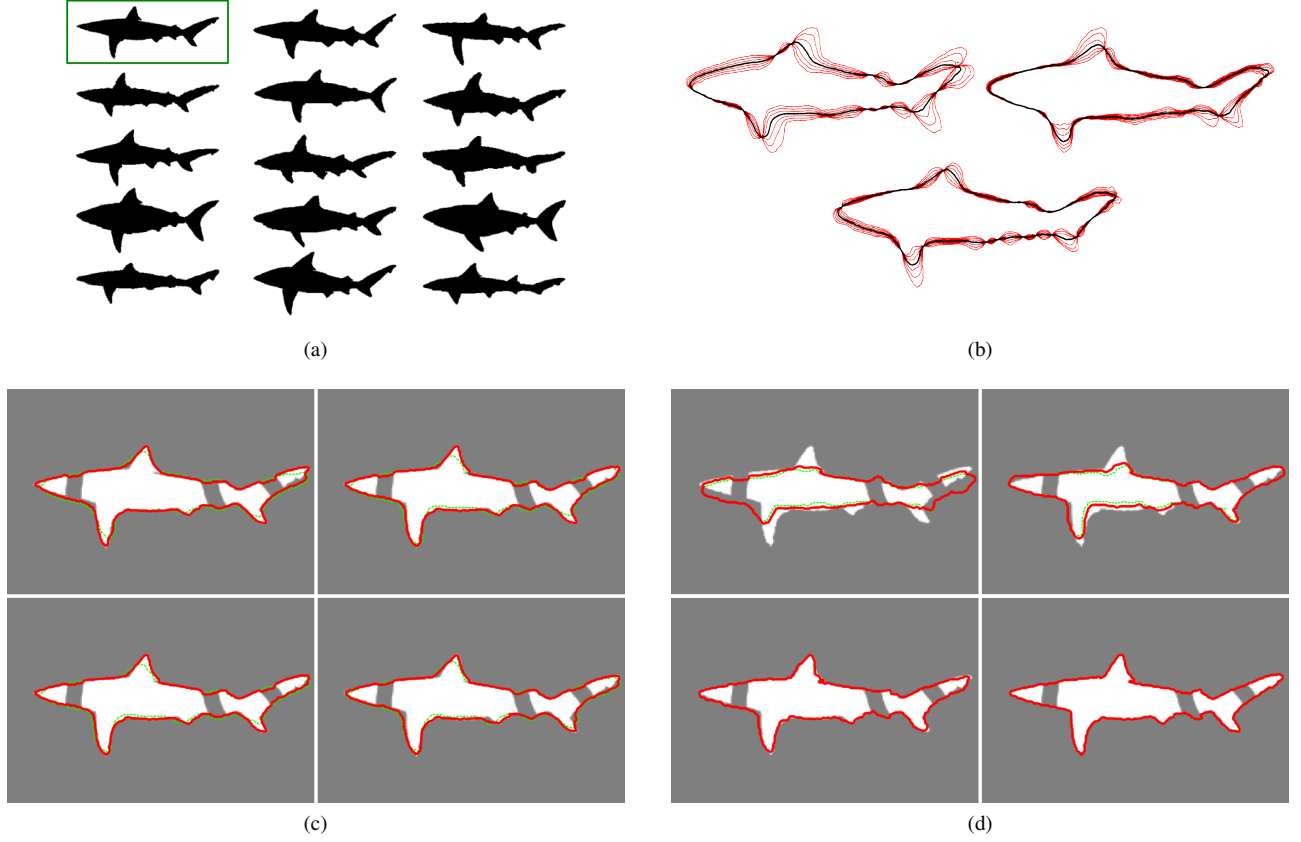
(a)

(b)

(c)

(d)

Fig. 6. Obtained segmentations with different orders of the shape model. (a) Set of shapes used in the experiment, the boxed shape is ϕ_S. (b) First three modes of variation of the model $\mathcal{M}_{S_1}$. (c) Segmentation of $\hat{\phi}_S$ (an occluded version of ϕ_S) when ϕ_S is not in the training set (left to right and top to bottom $d = [3, 7, 10, 13]$). (d) Segmentation of $\hat{\phi}_S$ when ϕ_S is in the training set (left to right and top to bottom $d = [3, 10, 13, 14]$). (This figure is in colors.)

dissimilarity between the shapes in the training set and the mean shape, $\hat{\Upsilon}_k$. This indicates the high accuracy of the proposed framework for this data.

Table II shows the dissimilarity measures between the obtained segmentations and the mean shapes for all the models in $\mathfrak{M}$, for the five different images and shapes in figures 4 and 5, the minimimum for each is in bold in each row, and obtained in the diagonal as expected from a correct model selection. Note how the difference between the minima and the next greater value in each row are considerable. This further indicates how the automatic model selection is correctly performed. Taking into account that the obtained segmentations correspond to shapes generated by the selected models, in all the experiments, the results were validated with the proposed measure, Equation (10), and were not validated by the other four used models. This further supports the validity of the proposed framework in general and the on-line automatic selection of the correct model in particular.

Figure 5b shows the obtained segmentations with the proposed framework for four different gray level images. The configuration of the framework is the same as in the previous examples, using the set of models $\mathfrak{M} = \{\mathcal{M}_{w_k}\}_{k=1}^5$. The automatically selected models, as well as the obtained segmentations, are also correct and accurate.

C. Varying the order of the models

This section further analyzes the segmentations when the order d of the model varies. The shapes used in this test are shown in Figure 6a. They are fifteen shapes of sharks taken from the SQUID database [19]. Two different sets of shapes are defined, Φ_{S_1} and Φ_{S_2}. Φ_{S_1} has fourteen shapes, leaving out the shape marked with a box in Figure 6a, while Φ_{S_2} uses the fifteen shapes. Two different models were created, $\mathcal{M}_{S_1}^d$ from Φ_{S_1} and $\mathcal{M}_{S_2}^d$ from Φ_{S_2}. Since Φ_{S_2} is larger than Φ_{S_1}, $\mathcal{M}_{S_2}^d$ might have more maximal modes of variation than $\mathcal{M}_{S_1}^d$, this happen in this case being $d = 14$ the number of modes of variation for $\mathcal{M}_{S_2}^d$ and $d = 13$ for $\mathcal{M}_{S_1}^d$. Figure 6b shows the first three modes of variation of $\mathcal{M}_{S_1}^d$. The modes of variation of $\mathcal{M}_{S_1}^d$ are very similar to those of $\mathcal{M}_{S_2}^d$.

Inspite of the "visual" similarity of the shapes in the set, their variations are larger than in the previous examples. For instance, they are not just a sampling of the deformation of an object like the walking sequence or the ellipses. This can be observed from the mean dissimilarity measure between the shapes in the training set and the mean shape, $\bar{\Upsilon}$. For the models $\mathcal{M}_{S_k}$ this value is close to 6 (see tables IIIa and IIIb) whereas for the models $\mathcal{M}_{w_k}$ is smaller than 4 (see Table I). This is a significant difference for this dissimilarity measure. (Note that the dissimilarities can be compared since they are normalized by their corresponding curve length.)

In order to analyze the segmentations varying the order of

the model, at first a single model is used in $\mathfrak{M}$ at a time, without the influence of the model selection component of the framework.

The first experiment consists of the segmentation of an input image with an occluded version of the boxed shape in Figure 6a, with the model $\mathcal{M}_{\mathrm{S}_1}^d$ with different order d (number of modes of variation). Let ϕ_{S} be the original shape, $\hat{\phi}_{\mathrm{S}}$ its occluded version, and $\tilde{\phi}_{\mathrm{S}_d}$ the obtained segmentation with the d-order model. Figure 6c shows the obtained segmentations $\tilde{\phi}_{\mathrm{S}_d}$ (red curves) for $d = [3, 7, 10, 13]$. The projection to the model is also plotted (dashed green curves).

This experiment is repeated using the model $\mathcal{M}_{\mathrm{S}_2}^d$. Figure 6d shows the obtained segmentations for $d = [3, 10, 13, 14]$ and the corresponding projections onto the model.

Table IIIa shows, for the model $\mathcal{M}_{\mathrm{S}_1}$, the dissimilarity measure between the obtained segmentations $\tilde{\phi}_{\mathrm{S}_d}$ for different d and the original non-occluded shape, $\Upsilon(\tilde{\phi}_{\mathrm{S}_d}, \phi_{\mathrm{S}})$, and the dissimilarity measure with respect to the mean shape μ_{S_1}, $\Upsilon(\tilde{\phi}_{\mathrm{S}_d}, \mu_{\mathrm{S}_1})$. Table IIIb shows the same results for the model $\mathcal{M}_{\mathrm{S}_2}^d$ and the mean shape μ_{S_2}.

As can be observed in both experiments, the projection better represents the shape as the order increases.

In the first experiment (Figure 6c), the obtained segmentation improves the adjustment to the shape as the model has more details to represent. This can be seen for example in the pectoral and tail fin and under the head. However, since the projection does not perfectly adjust to the object, there is a competition between both energy terms, generating an intermediate curve that does not completely fit the present object. If more weight is added to the E_{SM} term, other regions of the curve, in non-occluded areas of the object, will follow a less accurate approximation of the projection and lead to a worst segmentation, for example in the belly of the shark. Trying to choose the best segmentation from this four cases of d, the curve obtained with $d = 10$ seems to be the slightly more accurate than the curve obtained with $d = 13$, for example, analyzing the adjustment in the tail and the pectoral and pelvic fins. This is also support by the dissimilarity between the obtained segmentations $\tilde{\phi}_{\mathrm{S}_d}$ and the original non-occluded shape, last column of Table IIIa. This provides an example of a kind of over-fitting of the model to the shapes in the training set, capturing features too specific in the higher eigenmodes.

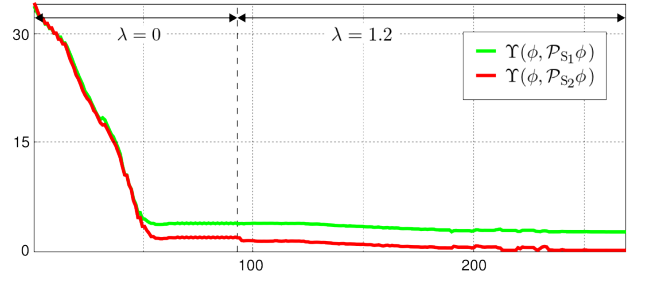On the other hand, with the model $\mathcal{M}_{\mathrm{S}_2}^d$ that includes the

Fig. 7. Evolution of the shape dissimilarity measure, $\Upsilon(\phi, \mathcal{P}_{\mathrm{S}_k}\phi)$, with the minimization iterations. (This figure is in colors.)

non-occluded shape in the training set (Figure 6c), the E_{SM} shape term has more relevance in the segmentation. When the order of the model increases, the projection gets more accurate and the segmentation improves. When the order is low and the projection is not accurate, the segmentation again is a compromise between the two energy terms, being an intermediate curve between the projection and the edges of the gray level information. Finally, the main difference between the segmentations with $d = 13$ and $d = 14$ are the fine details like high curvature points, see the extreme points in the tail and the pectoral fin.

The last experiment of this section is done using both models in the set $\mathfrak{M} = \{\mathcal{M}_{\mathrm{S}_1}^{13}, \mathcal{M}_{\mathrm{S}_2}^{14}\}$, and the same occluded shape in the input image. Figure 7 plots the dissimilarity measure for this example. The selected model is $\mathcal{M}_{\mathrm{S}_2}^{14}$ which has one additional eigenmode and obtains a better description. The obtained segmentation is, of course, the same segmentation shown in Figure 6d with $d = 14$. This further supports the necessity of high-order models in order to obtain accurate segmentations, in particular when the different object classes are relatively similar.

These obtained segmentations are also validated by the proposed validation process, (Equation (10)).

V. INVARIANCE TO TRANSLATION

Invariance to geometric transformations (such as translations, rotations and scaling) is a desiderable property in a general framework for segmentation. One way to do this is to substitute $H(\phi)$ by

$$H(\sigma R_\theta(\phi(x - x_0)))$$

in Equation (7), as in the work of Cremers et al. [2, section 5.1]. Here, σ is a scale factor, R_θ a rotation matrix of a given angle θ, and x_0 a translation vector.

This section proposes an extension of the functional in Equation (7), adding invariance to translation (other invariance are similarly added).

Consider all the shapes aligned with respect to their corresponding center of mass p_ϕ, defined for a certain shape ϕ as

$$p_\phi = \frac{\int_\Omega p H(\phi(p)) \mathrm{d}p}{\int_\Omega H(\phi(p)) \mathrm{d}p}.$$

The shape models are build in the same way as in the previous section. Considering that all the shapes in Φ_k have the same

$\mathcal{M}_{\mathrm{S}_1}^d$			$\mathcal{M}_{\mathrm{S}_2}^d$		
d	$\Upsilon(\tilde{\phi}_{\mathrm{S}_d}, \mu_{\mathrm{S}_1})$	$\Upsilon(\tilde{\phi}_{\mathrm{S}_d}, \phi_{\mathrm{S}})$	d	$\Upsilon(\tilde{\phi}_{\mathrm{S}_d}, \mu_{\mathrm{S}_2})$	$\Upsilon(\tilde{\phi}_{\mathrm{S}_d}, \phi_{\mathrm{S}})$
3	3.59	1.46	3	5.68	6.34
7	4.15	0.97	10	5.49	2.60
10	4.25	0.95	13	3.69	1.63
13	3.96	1.38	14	4.00	0.63
(a) $\bar{\Upsilon} = 6.11$, $\sigma_\Upsilon^2 = 1.42$			(b) $\bar{\Upsilon} = 5.93$, $\sigma_\Upsilon^2 = 1.48$		

TABLE III
NUMERICAL VALIDATION RESULTS (EQUATION (10)) AND DISSIMILARITY MEASURES (EQUATION (8)) FOR THE OBTAINED SEGMENTATIONS $\tilde{\phi}_{\mathrm{S}_d}$ SHOWN IN FIGURE 6.
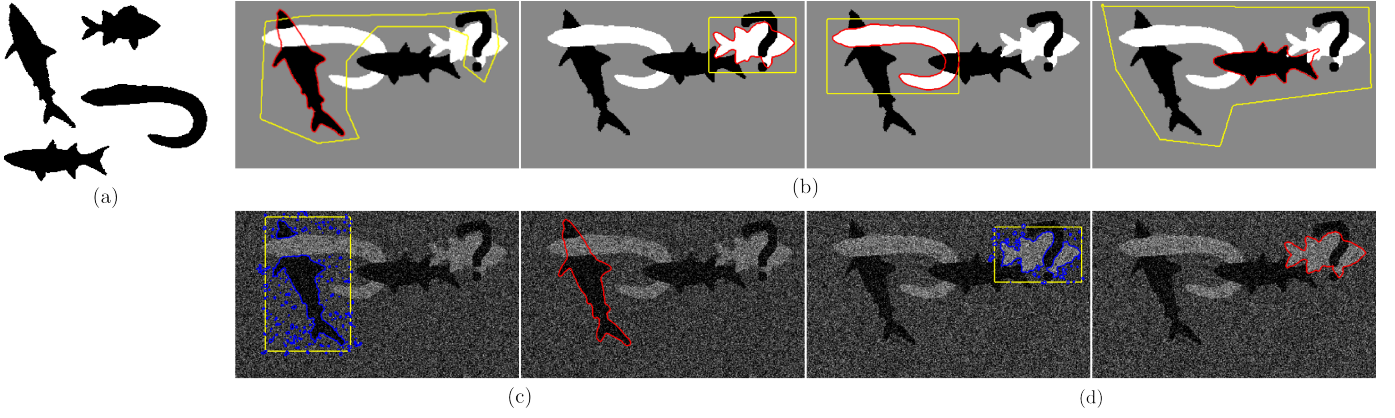
Fig. 8. Segmentations obtained with the translation invariant energy (see Equation (13)). (a) Four zero-order models (shape prior templates). (b) Four different initial curves (yellow curves) and the obtained segmentations (red curves). (c)-(d) Two different initializations with Gaussian noise added. Initial curve (in yellow), last curve previous to the "prior activation" (in blue) and obtained accurate and valid segmentation (red curve).

center of the mass p_k, this point becomes the center of mass of the model. To obtain the projection of a new, not aligned, shape ϕ with center of mass p_ϕ, first the shape is translated to p_k and then projected. Mathematically, the projection to the translation invariant model $\mathcal{M}_k^{(\text{IT})}$ centered at p_k, becomes $\mathcal{P}_k\phi(p + p_\phi - p_k)$, and lets call p_ϕ^k its center of mass. Finally, the projection needs to be translated back to the original center of mass. Defining $T(\phi(p), p_0) = \phi(p + p_0)$, the final projection to the translation invariant model is

$$\mathcal{P}_k^{(\text{IT})}\phi = T(\mathcal{P}_k T(\phi(p), p_\phi - p_k), p_\phi^k - p_\phi). \quad (12)$$

Without loose of generality $p_k = 0$ is assumed from now on.

In order to incorporate the invariance to translation in the original energy, the shape models terms $E_k(\phi, \mathcal{M}_k)$ become

$$E_k(\phi, \mathcal{M}_k) = \int_\Omega \|H(\phi(p + p_\phi)) - H(\mathcal{P}_k\phi(p + p_\phi))\|^2 \, \mathrm{d}p. \quad (13)$$

If the kth model provides a good representation of ϕ, the corresponding centers of mass are close, $p_\phi^k \approx p_k = 0$. For the derivation of the corresponding gradient descent expression below, this approximation is assumed, simplifying the deduction here presented. However in the implementation the actual p_ϕ^k is used. The updated gradient descent expression is given by

$$\frac{\partial E_{\text{SP}}}{\partial \phi} = -2\sum_{k=1}^M \beta_k \Big[\Delta H(p)\Delta\delta(p) + $$
$$\frac{\delta(\phi(p))(p - p_\phi)^T}{\int_\Omega H(\phi(u))\mathrm{d}u} \int_\Omega \Delta H(z)\Delta\delta(z)\nabla\phi(z)\,\mathrm{d}z \Big], \quad (14)$$

where $\Delta H(p) = \left(H(\phi(p)) - H\left(\mathcal{P}_k^{(\text{IT})}\phi\right)\right)$ and $\Delta\delta(p) = (\delta(\phi(p)) - \delta(T(W\mathcal{P}_k T(\phi(p), p_\phi), -p_\phi)))$.

Again, β_k is treated as static, as a first order approximation for the gradient descent.

A. Model selection with invariance to translation

An example of the model selection capabilities of the translation invariant framework is shown in Figure 8. In order to test only the model selection, without being influenced by the

adjustment to the selected model, four zero-order models are used. A zero-order model obtains always the same synthesized shape for any input shape, this synthesized shape being the mean shape of the model in this work. The zero-order models are shown in Figure 8a. These four different shapes from the SQUID database are arranged in a single image with occlusions for each shape and this becomes the input image for testing the framework. Figure 8b shows four different initial curves (in yellow) and the segmentation (red curves) obtained with the proposed framework. Two examples with Gaussian noise added to the image are shown in Figures 8c and 8d.

In all the cases the segmentations are accurate which also implies that the selected model is the correct one. Note that these results are valid even when the initial curves are not clearly defining one object (following our definition of validity, Equation (10)).

B. Segmentations with invariance to translation

The last tests show the translation invariant framework working with the high-order models. The first experiment reproduces the test with the ellipses (Section IV-A), now with the addition of Gaussian noise and the translation of the ellipse. Figure 9a shows the details of the segmentation. The first subfigure shows the initial curve (in yellow), an intermediate step (blue curve), and its projections to both models (green and red dashed curves). Note the projections translated to the center of mass of ϕ. The second subfigure shows the projection to the correct model and the curve filling the occlusions. The last subfigure shows the obtained segmentation (red curve).

Figure 9b shows the segmentation of an image with an occluded binary shape from the walking sequence and Figure 9c shows the segmentation of an image with a gray level shape from the walking sequence, using the set of translation invariant models for this dataset. Again the result is an accurate segmentation with a valid shape from the correctly selected model.
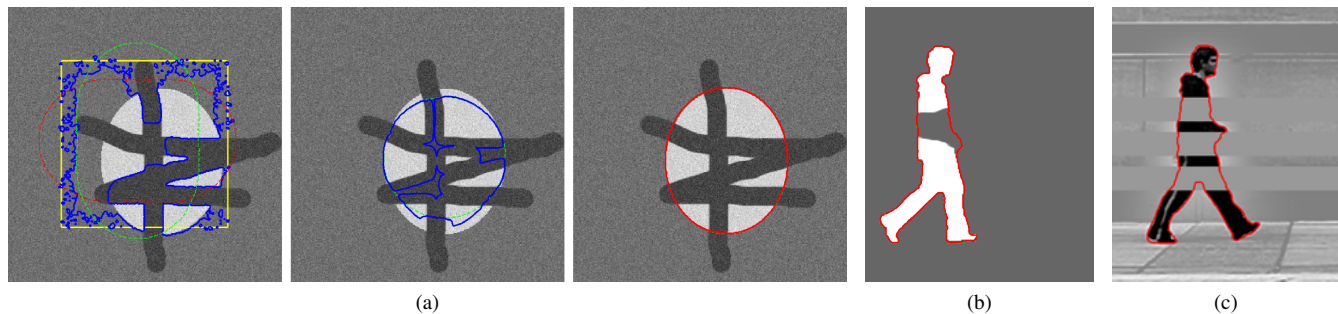
Fig. 9. Segmentations obtained with the translation invariant energy (see Equation (13)). (a) Three steps in the segmentation of the ellipse from Figure 3. Note in the first image the projections translated to the center of mass of ϕ. The second image shows the projection to the correct model and the curve filling the occlusions. The last image shows the obtained segmentation (red curve). (b) Obtained segmentation of a binary shape from the walking person cycle. (c) Obtained segmentation of a gray-valued shape from the walking person cycle. (This figure is in colors.)

VI. Concluding remarks

A framework for simultaneous and automatic model selection and object segmentation was introduced in this work. The proposed technique is based on a new energy that combines region based segmentation with on-line selection of the best model for the object present in the image, and an adjustment to the best description of the object given the selected model.

The segmentation is obtained via gradient descent energy minimization, and the model selection is automatic in each iteration, without the need to run the segmentation with all the models and then select the best solution. The on-line decision of best description is based on a shape dissimilarity measure between the curves. The selection is such that a unique model candidate is considered at each step of the minimization. Invariance to shape transformations are incorporated into the proposed framework as well.

Possible directions for further improvements include incorporating high-order modes in the validation step and considering going beyond PCA, as well as including class-dependent model orders (d_k). Results in these directions will be reported elsewhere.

Acknowledgments

References

[1] M. E. Leventon, W. E. L. Grimson, and O. D. Faugeras, "Statistical shape influence in geodesic active contours," in *CVPR*, vol. 1, 2000, pp. 1316–1323. [Online]. Available: http://ieeexplore.ieee.org:80/xpls/abs_all.jsp?isNumber=18553&prod=CNF&arnumber=855835&arSt=+316&ared=+323&arNumber=855835

[2] D. Cremers, S. Osher, and S. Soatto, "Kernel density estimation and intrinsic alignment for shape priors in level set segmentation," *IJCV*, vol. 69, no. 3, pp. 335–351, 2006. [Online]. Available: http://dx.doi.org/10.1007/s11263-006-7533-5

[3] A. Tsai, A. J. Yezzi, W. M. W. III, C. Tempany, D. Tucker, A. C. Fan, W. E. L. Grimson, and A. S. Willsky, "Model-based curve evolution technique for image segmentation," in *CVPR*, vol. 1, 2001, pp. 463–468. [Online]. Available: http://ieeexplore.ieee.org:80/xpls/abs_all.jsp?isNumber=21353&prod=CNF&arnumber=990511&arSt=+463&ared=+468&arNumber=990511

[4] G. Charpiat, O. D. Faugeras, and R. Keriven, "Shape statistics for image segmentation with prior," in *CVPR*, 2007, pp. 1–6. [Online]. Available: http://dx.doi.org/10.1109/CVPR.2007.383009

[5] D. Cremers, N. A. Sochen, and C. Schnörr, "A multiphase dynamic labeling model for variational recognition-driven image segmentation," *IJCV*, vol. 66, no. 1, pp. 67–81, 2006. [Online]. Available: http://dx.doi.org/10.1007/s11263-005-3676-z

[6] N. Vu and B. S. Manjunath, "Shape prior segmentation of multiple objects with graph cuts," in *CVPR*, 2008, pp. 1–8. [Online]. Available: http://dx.doi.org/10.1109/CVPR.2008.4587450

[7] D. Cremers, T. Kohlberger, and C. Schnorr, "Shape statistics in kernel space for variational image segmentation," *Pattern Recognition*, vol. 36, no. 9, pp. 1929–1943, sep 2003. [Online]. Available: http://www.sciencedirect.com/science/article/B6V14-48DXN97-9/2/11741a8017ce4cebec21ae0524796a7a

[8] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models-their training and application," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38–59, 1995. [Online]. Available: http://dx.doi.org/10.1006/cviu.1995.1004

[9] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266–277, 2001. [Online]. Available: http://dx.doi.org/10.1109/83.902291

[10] P. Arias, G. Randall, and G. Sapiro, "Connecting the out-of-sample and pre-image problems in kernel methods," in *CVPR*, 2007, pp. 1–8. [Online]. Available: http://dx.doi.org/10.1109/CVPR.2007.383038

[11] R. R. Coifman and S. Lafon, "Diffusion maps," *Applied and Computational Harmonic Analysis*, vol. 21, no. 1, pp. 5–30, 2006. [Online]. Available: http://dx.doi.org/10.1016/j.acha.2006.04.006

[12] U. von Luxburg, "A tutorial on spectral clustering," *Statistics and Computing*, vol. 17, no. 4, pp. 395–416, 2007. [Online]. Available: http://dx.doi.org/10.1007/s11222-007-9033-z

[13] B. Schölkopf and A. Smola, *Learning with Kernels*. Cambridge, MA, USA: MIT Press, 2002. [Online]. Available: http://www.learning-with-kernels.org

[14] S. Dambreville, Y. Rathi, and A. Tannenbaum, "Shape-based approach to robust image segmentation using kernel pca," in *CVPR*, vol. 1, 2006, pp. 977–984. [Online]. Available: http://dx.doi.org/10.1109/CVPR.2006.279

[15] N. Paragios and R. Deriche, "Geodesic active regions for supervised texture segmentation," in *ICCV*, 1999, pp. 926–932. [Online]. Available: http://dx.doi.org/10.1109/ICCV.1999.790347

[16] G. Sapiro, R. Kimmel, and V. Caselles, "Geodesic active contours," in *ICCV*, 1995, pp. 694–699. [Online]. Available: http://dx.doi.org/10.1109/ICCV.1995.466871

[17] T. Funkhouser, M. Kazhdan, P. Shilane, P. Min, W.Kiefer, A. Tal, S. Rusinkiewicz, and D. Dobkin, "Modeling by example," *ACM Transactions on Graphics*, vol. 23, no. 3, pp. 652–663, Aug. 2004.

[18] G. Kanisza, *Organization in Vision*. New York, NY: Praeger, 1979.

[19] S. Abbasi, F. Mokhtarian, and J. Kittler, "Squid database (shape queries using image databases)," WWW, 1996, visited, April 26th. 2008. [Online]. Available: http://www.ee.surrey.ac.uk/CVSSP/demos/css/demo.html